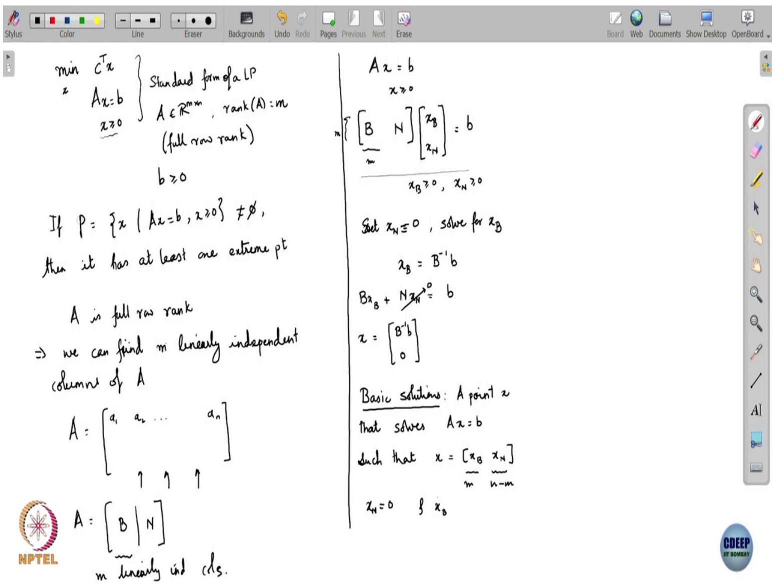
Optimization from Fundamentals
Prof. Ankur Kulkarni
Department of Systems and Control Engineering
Indian Institute of Technology, Bombay

Lecture - 11A
Equivalence of extreme point and BFS

So, in the previous class we ended with this observation that you can that any linear program can be written in this form, that you are minimizing a function c transpose x over constraints that look like this.

(Refer Slide Time: 00:29)



So, a linear program is an optimization of a linear function over a polyhedron, but then you can write it in this form that you are minimizing a linear function c transpose x over a set of the form A x equal to b and x greater than equal to 0. This is what was called the standard form of a linear program.

And what is what are the assumptions of the standard form? So, if A is a matrix that is R m cross n, then the rank of A should be m. So, A is full row rank and moreover we said we can assume b to be greater and moreover we assume that b could be greater has to be greater than equal to 0, alright.

So, now what I mentioned briefly at the end of the last class was that; such of an l p once written in a standard form. If it has a if it has a non empty feasible region, if this feasible region let us call this let us call this P this polyhedron, x such that A x equals b and x greater than equal to 0.

If this is not empty right, if it is not empty then it has at least one extreme point. And we also said showed how that extreme point is derived it is derived by looking at certain columns of A that form a linearly independent set, ok. So, what I will do today is just whatever I said in the previous class, I will just do it in a little bit more detail, ok.

So, now, A remember is A is full row rank, which means that we can find we can find m linearly independent columns of A, we can find m linearly independent columns of A, ok. So, now can we find more than m linearly independent columns of A can we have more than m linearly independent columns, that is not possible because A the rank of A is m, right.

So, we can find, but we can certainly find m linearly independent columns. Now, let us put let us look at A like this. So suppose A is this sort of matrix suppose this is a 1, this is a 2 this is a n. Now, I can find say m linearly independent columns like this, ok. Now, they may not be in they may not be the first m columns or the second, or they may not be contiguous one after the other, but there are some m linearly independent columns.

Now, what I will do is for simplicity, I will just arrange them in this sort of form I will just put take all the linearly independent columns and write them first, ok. So, I express A in this sort of form, I write A as B N, where B comprises of linearly independent say m linearly independent columns.

Now, what this amounts to doing just putting B at the beginning and all the rest at the end, what it amounts to doing is simply permuting the order of my x's, right. The ones that have the x's that are that correspond to the columns the linearly independent columns have just been written first and the rest have been written later, right. And that does not affect my other constraint, which is just x greater than equal to 0.

So, without loss of generality for each selection of linearly independent columns I can express A in this kind of form by requiring maybe if necessary permuting x, ok. So, what I so, remember I need to find if I need to be feasible with respect to my constraints. So, which means that I still need to satisfy A x equals b and x greater than equal to 0.

Now, I have B N, which is my A. Let us suppose the columns that multiply the x's that multiply b are written as x B and the rest are written as x N. So, then I am looking to satisfy this equation B N x B x N equals b, in addition to that I am I need x B greater than equal to 0 and x N also greater than equal to 0.

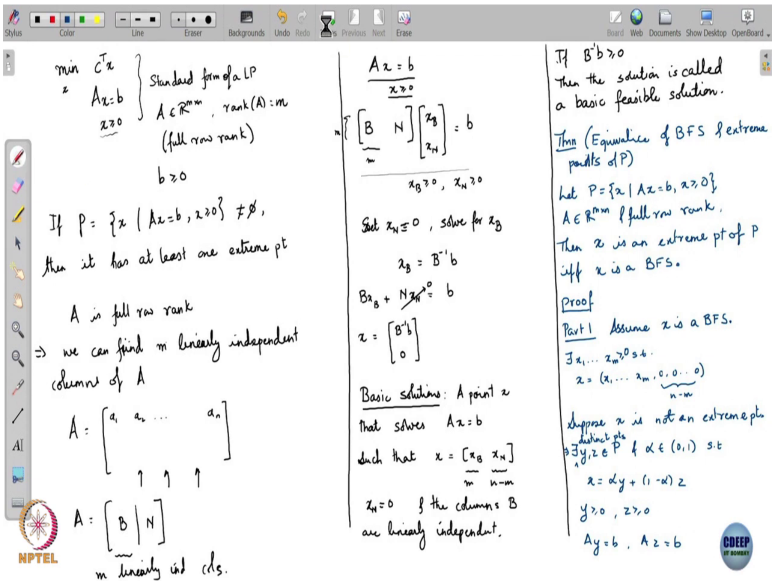
Now, because B has m rows and m columns so, B is square and is has is non singular because it is rank is m, it is non singular. So, what that means is; I can one way of solving this equation here is to do the following. I just set x N equal to 0, set x N as 0 and then solve for B N, sorry solve for x B in that case then what would I get x B as I get x B as capital B inverse b, right.

How do I get this? This is because this is B x B plus N x N equals small b and I am setting x N equal to 0. So, then in that case I get that x B is B inverse b. So, in summary the point x is of the form B inverse b followed by 0.

Once, I have done this actually it did not you realize that, it did not really matter that I put my columns of b first and the you know the columns of b first and the rest later. What matters is so long as I am keeping track of what columns form b. I can still do this right, so what this means is I can create points like this these are called basic solutions, ok.

Basic solutions: basic solution is point of the where is obtained is a point x that solves A x equal to b; such that x can be written as x B x N. Now, when I write it like this it does not necessarily mean that they are in this order, but they are basically there are n there are n minus m here and m here x N equals 0 and x B.

(Refer Slide Time: 08:56)



And the columns of B are linearly independent. So, to generate basic solutions all we need to do is look at some m linearly independent columns of A. Corresponding to those columns you will solve for x b all the other x a x a all the other coordinates of x are set as 0, right, put the two together that gives you what is called a basic solution.

Now, such a basic solution need not necessarily satisfy x greater than equal to 0. So, remember we had to satisfy this as well as this. So, a basic solution definitely satisfies the first equation here A x equals b, but it need not satisfy x greater than equal to 0. So, if it so, happens that B inverse b is greater than equal to 0, then it will satisfy A x equal x greater than equal to 0.

So, if B inverse b is greater than equal to 0 then the solution is called a basic feasible solution, then the solution is called a basic feasible solution alright. Now, what is the why is this concept so, important?

The reason it is so, important is because basic feasible solutions, which is a way of solving these linear equations and this greater than equal to inequality. Basic feasible solutions, which are just basically an algebraic solution of this system of equations and inequalities; they actually have a geometric significance, ok.

Basic feasible solutions have this geometric significance and that is given by this theorem, ok. This what does it talk about it talks of the equivalence of basic feasible solutions and extreme points, ok. So, let P be the set, let P be the set x; such that A x equals b x greater than equal to 0, A in is in R m cross n and full row rank then x is an extreme point of P, if and only if x is a basic feasible solution, right.

So, let us see how this is proved so, what does this theorem say it is saying that the there is a fundamental equivalence between the notion of extreme points, which is a geometric property of the feasible region of this of a linear program. And this algebraic construct called as a basic feasible solution ok.

So, what is the, what is the proof it is this the proof is not that hard. So, you will see that the proof actually uses the fact that A is full row rank ok that is vital in all of this ok. So let us see we have to prove that x is the extreme point of P, if and only if x is a basic feasible solution, alright. So, now, let us assume let us first do this direction suppose we let us prove that x is a basic feasible solution and show that it must be an extreme point.

And then we will as then we will show the other direction which is that which is showing that if x is an x is an extreme point then it must be a basic feasible solution, alright. So, part 1: so, assume x is a BFS ok, assume x is a basic feasible solution. What does this mean? This means there exist some x 1, x 2 dot, dot, dot x m, which are the which are; such that which are coordinates of x; such that x can be written like this x 1, x 2, till x m and the rest of it are 0.

So, now the x 1 till x m also could be 0 that does not mean that they are not 0, but definitely the remaining n minus m have to be 0 right, but x 1, till x m are necessarily greater than equal to 0, because this is the base x is a basic feasible solution, right. So, let us write it like this there exists x 1, x 2, x m greater than equal to 0; such that x can be written in this form x 1 till x m followed by a string of n minus m 0's, we assume that x is a BFS.

Now, suppose x is not an extreme point suppose x is not an extreme point if x is not an extreme point. Then there exists this implies there exists some y and z in this set P and an alpha that is strictly between 0 and 1, y and z.

There exists distinct points y and z, distinct points y and z in P and an alpha in 0 1; such that x can be written as alpha y plus 1 minus alpha z right, now x can be written as alpha; so, x is a convex combination of y and z, y and z are distinct points in P and alpha is something that strictly between 0 and 1.

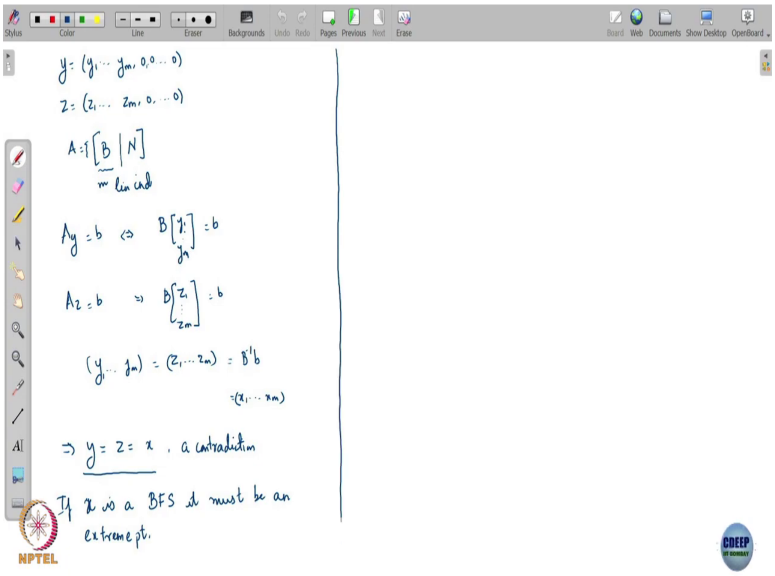
Now, if that is the case what can one say about what can you say about the components of y and z. If you look at x; x has this sort of form it is the first x 1 till x m the first m, m of these are some elements greater than equal to 0, but definitely the last n minus m are 0.

Now, the if the last n minus m are 0, then that has to be the case also for the last n minus m components of y as well as z right. Because it is a convex combination of these, which is giving you; which is giving you; which is giving you x, right. So, formally see remember y is greater than equal to 0, z is also greater than equal to 0.

Why is this is the case? Because y and z both belong to P. If they since they belong to P, y and z are also greater than equal to 0, they also satisfy A y equals b and A z equals b. Now, if you look at the last n minus m components, if you look at the last n minus m components in this equation you are getting a 0 here on the left hand side.

The only and that is coming from a convex combination of non negative terms of y and non negative terms from z, the only way that you can get they can add up to 0 is they are both themselves 0, right. So, A by the last n minus m components of y and of z have to be 0, right.

(Refer Slide Time: 18:35)



So, what that means, is y is of the form y 1 till y m followed by 0, 0, 0, z is of the form z 1 till z m followed by 0, 0, 0, ok alright. Now, therefore now go back to this equation, which is that A y equals b we have already that A y equals b and A z equals b, right. If A y is equal to b and A z is equal to b. Now, suppose I had partitioned A as B N, where these are my these are all linearly independent. Then A y equals b is simply is simply saying is equivalent to saying B into y 1 till y m equals small b.

And similarly A z equals b is saying B into z 1 till z m equals small b, right. But then but then what do you find well, b is b has linearly independent columns m b is non singular, b is invertible. So, which means I should be, I can just take b on the others I can invert and take b on the other side and that will give me that y 1 till y m, this is equal to z 1 till z m and they are both equal to B inverse b.

But then what was B inverse b it was actually x 1 till x n. So, what we conclude from here is that you must have y equal to z equal to x. So it so, this from this we get that, basically that if x is a basic feasible solution it must be an extreme point.

So, y equals z equals x this is a contradiction, why is this is a contradiction? This is a contradiction, because we had assumed that because we had assumed that y and z are distinct points in P, for x to not be an extreme point this is how it should be y and z must be distinct points, ok. So, x is; so, x is if x is a basic feasible solution it must be an extreme point, alright.